Data Dictionary and Background

The dataset is provided by MobData, Inc. The vendor’s main value add is providing location data for a large sample of mobile devices linked to individuals. This data is collected via pings to the MobData network. Through these pings, the company triangulates the individual’s location in longitude and latitude. For the current application, this location information is used to identify whether the individual was in the vicinity of a home improvement store. This identification is done using proprietary software of the company. Ultimately, MobData provides data with visit information that links users to venues. In some products, they also provide other geocodes for the locations.

MobData provides four different data tables. The first is a list of users who visited at least one home improvement store. This table contains location information for the user. The second is a list of the chains. The third is a list of venues classified by chain that gives location information for the venues. The final table is a list of visits to venues by users.

The data comes from a full year of mobile location data. The data is already filtered to visitors of the five major chains of home improvement stores. The included set of users is intended to represent the population of shoppers. *You should proceed as if the sample is a legitimate representative sample.*

On the next page, is more information about the specific attributes for each of the tables.

*NOTE: This data is simulated! The real data contains more variables (e.g., date and time at venue). Due to vendor restrictions, the case data is simulated. However, the data is simulated to have similar broad features to the real data. For the purposes of the assignment it is as good as the real data. However, do NOT try to use this data for real business decisions, since the data is simulated!*

users

* 1. UserID – Numeric identifier
  2. Latitude – Latitude of home in decimal degrees (estimated)
  3. Longitude – Longitude of home in decimal degrees (estimated)

chains

1. chain – name of chain corresponding to venueName

venues

* 1. venueID – Numeric identifier
  2. venueName – Text name of chain for this venue (MobData proprietary places data)
  3. latitude – Latitude in decimal degrees for venue (MobData proprietary places data)
  4. longitude – Longitude decimal degrees for venue (MobData proprietary places data)

visits

* 1. UserID
  2. venueID